On the Cause of Hierarchical Order among the Correlation Coefficients of a Number of Variates Taken in Pairs.

By Godfrey H. Thomson, D.Sc., Armstrong College, in the University of Durham.

(Communicated by Prof. T. H. Havelock, F.R.S. Received January 2, 1919.)

CONTENTS.

	PAGE
Introduction	400
An Experiment in Chance	401
Hierarchical Order Produced by Sampling the Population	405
Hierarchical Order Produced by Sampling the Elements which make up the Variates	407

Introduction.

The object of this paper is to investigate the cause of a phenomenon in the realm of chance, which has become technically known as hierarchical order among correlation coefficients, and has been held to prove the existence of a general factor running through the correlated varieties, and the absence of group factors running through some but not through all of them. The question arose in the science of experimental psychology, but it is here, after the introductory paragraphs, considered as a general question in probability.

When mental tests are applied to a number of subjects, and the correlations between the marks are calculated for every possible pair of tests, the correlation coefficients obtained show, as a rule, a tendency to arrange themselves in hierarchical order. By this is meant that the order of sequence of the mental tests, according to the size of the correlation of each with a fixed one of their number, proves to be largely independent of the choice of this latter. If the mental tests, which we may call by the names x_1, x_2, x_3, ..., have been arranged in order according to the total correlation of each with all the others, and if a square table such as the following be formed :—

	x_1	x_2	x_3	x_4	......
x_1	.	r_{12}	r_{13}	r_{14}	......
x_2	r_{21}	.	r_{23}	r_{24}	......
x_3	r_{31}	r_{32}	.	r_{34}	......
x_4	r_{41}	r_{42}	r_{43}	.	......
...	......	......	......	......	......

then the hierarchical order shows itself in the fact that there is a tendency for each correlation coefficient r to be greater than its neighbour on the right and its neighbour below it. A method, which has been used for measuring the degree of perfection of the hierarchical order, is to take the correlation of each pair of columns of the above table. Clearly, if hierarchical order is present, all these correlations will be high, and in the most perfect case will become unity.

From the fact that there is a strong tendency toward hierarchical order among correlation coefficients in mental tests (though not so strong as certain methods of estimating it would suggest),* the conclusion has been drawn that all these correlations are due to the presence of a general factor running through all the tests, and that none of them are due to group factors which run through two or more tests but not through all. Now, while there is no doubt but that such a general factor, combined with such an absence of group factors, would produce hierarchical order, it did not seem to be proved that such order could not be produced in other ways, if not to the highest degree, yet, at any rate, to a very considerable degree of perfection, sufficient to explain the experimental results. I therefore endeavoured to make artificial variates with dice throws, in a way which will be clear in the sequel, and soon succeeded in making hierarchical order (which passed Prof. Spearman's criterion for perfection in a hierarchy) without using any general factor, group factors alone producing the correlations.† I did not at first appreciate how much might be left to chance in this operation without seriously endangering the formation of a hierarchy, but later I found that all that was necessary was to postulate group and specific factors, leaving the number and identity of those in each variate to chance, viz., to the draw of a card from a pack. In other words, I came upon the following experiment.

An Experiment in Chance.

Write down the letters x_1, x_2, x_3, ... as the names of the variates to be formed, and prepare columns to receive the numbers of group factors and specific factors in each variate. Determine each number by the draw of a card from an ordinary playing pack, returning the card and shuffling between each draw: the knave, queen, and king may be counted 11, 12, and 13 respectively. The result of one such set of drawings is shown in this Table :—

---

* I propose to justify this statement in a separate paper.

† 'British Journal of Psychology,' vol. 8, p. 271 (1916).

Variate.	Group factors.	Specific factors.	Total.
x_1	5	5	10
x_2	5	3	8
x_3	12	12	24
x_4	1	3	4
x_5	7	6	13
x_6	9	5	14
x_7	13	13	26
x_8	1	2	3
x_9	9	3	12
x_{10}	11	5	16

Proceed next to identify the group factors of each variate. Do this by using a single suit of the pack. After shuffling it well, lay out the top five cards to represent the five group factors in x_1, and note them. After replacing them and reshuffling, do the same for x_2, and so on, as in this Table:—

	Ace	2	3	4	5	6	7	8	9	10	Kn	Q	K
x_1	/	/			/		/						/
x_2					/		/		/	/		/	
x_3	/	/	/	/	/		/	/	/	/	/	/	/
x_4												/	
x_5	/	/			/		/	/				/	/
x_6				/	/	/	/	/	/	/	/	/	
x_7	/	/	/	/	/	/	/	/	/	/	/	/	/
x_8		/											
x_9	/	/	/	/	/	/	/	/					/
x_{10}		/	/	/	/	/	/	/		/	/	/	/

The next step is to note the number of factors common to each pair of variates, as in this Table:—

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
x_1	—	2	5	0	5	2	5	1	5	4
x_2	2	—	5	1	3	5	5	0	2	4
x_3	5	5	—	1	7	8	12	1	8	10
x_4	0	1	1	—	1	1	1	0	0	1
x_5	5	3	7	1	—	4	7	1	5	6
x_6	2	5	8	1	4	—	9	0	5	8
x_7	5	5	12	1	7	9	—	1	9	11
x_8	1	0	1	0	1	0	1	—	1	1
x_9	5	2	8	0	5	5	9	1	—	8
x_{10}	4	4	10	1	6	8	11	1	8	—

From these, and from the total number of factors both specific and group in each variate, can be found the correlation which would occur between the variates were we to throw dice, one to each factor, and repeat the throwings a large number of times. The formula is

$$r = \frac{\text{Number of common factors}}{\text{Geometrical mean of totals}}.$$

This formula is applicable not only to variates formed by the addition of dice, but to variates which are any function of the factors or elements, provided that the form of the function is the same in each variate, and that the standard deviation is the same for each element or factor.* We thus obtain the following Table of theoretical correlation coefficients :—

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	Totals.
x_1	—	0·22	0·32	0·00	0·44	0·17	0·31	0·18	0·46	0·32	2·42
x_2	0·22	—	0·39	0·18	0·29	0·47	0·35	0·00	0·20	0·35	2·45
x_3	0·32	0·39	—	0·10	0·39	0·44	0·48	0·12	0·62	0·51	3·37
x_4	0·00	0·18	0·10	—	0·14	0·13	0·10	0·00	0·00	0·12	0·77
x_5	0·44	0·29	0·39	0·14	—	0·30	0·38	0·16	0·40	0·41	2·91
x_6	0·17	0·47	0·44	0·13	0·30	—	0·47	0·00	0·38	0·53	2·89
x_7	0·31	0·35	0·48	0·10	0·38	0·47	—	0·11	0·51	0·54	3·25
x_8	0·18	0·00	0·12	0·00	0·16	0·00	0·11	—	0·17	0·15	0·89
x_9	0·46	0·20	0·62	0·00	0·40	0·38	0·51	0·17	—	0·58	3·32
x_{10}	0·32	0·35	0·51	0·12	0·41	0·53	0·54	0·15	0·58	—	3·51

If we wished to obtain experimental values of these, dice would have to be thrown, one to each factor. The die corresponding to the group factor called "Ace" would have its score counted into every variate containing the group factor in question. The dice representing the specific factors would, of course, only be counted into the one variate in which they occur.

The last column of the preceding Table gives the total correlation of each variate with all the others, found by adding the rows of the square Table. Rearrange now the sequence of the variates in the order of magnitude of these totals, and we obtain the following Table :—

	x_{10}	x_3	x_9	x_7	x_5	x_6	x_2	x_1	x_8	x_4
x_{10}	—	0·51	0·58	0·54	0·41	0·53	0·35	0·32	0·15	0·12
x_3	0·51	—	0·62	0·48	0·39	0·44	0·39	0·32	0·12	0·10
x_9	0·58	0·62	—	0·51	0·40	0·38	0·20	0·46	0·17	0·00
x_7	0·54	0·48	0·51	—	0·38	0·47	0·35	0·31	0·11	0·10
x_5	0·41	0·39	0·40	0·38	—	0·30	0·29	0·44	0·16	0·14
x_6	0·53	0·44	0·38	0·47	0·30	—	0·47	0·17	0·00	0·13
x_2	0·35	0·39	0·20	0·35	0·29	0·47	—	0·22	0·00	0·18
x_1	0·32	0·32	0·46	0·31	0·44	0·17	0·22	—	0·18	0·00
x_8	0·15	0·12	0·17	0·11	0·16	0·00	0·00	0·18	—	0·00
x_4	0·12	0·10	0·00	0·10	0·14	0·13	0·18	0·00	0·00	—

---

* G. H. Thomson, *loc. cit.*, p. 275. It can readily be deduced from Bravais, 'Memoires de l'Institut de France,' 1846, IX, eqn. 28.

Here the tendency to hierarchical order is quite noticeable. This particular example is chosen from among a number which I have calculated, and I have purposely chosen that which shows the *least* hierarchical tendency. There is, however, clearly a general lowering of the coefficients as we pass either along a row or down a column. The columnar correlation is high for the first few columns. For the columns headed x_{10} and x_3 it is 0·97, and as far down as the columns headed x_5 and x_6 it is still 0·65. If these theoretical numbers were blurred by experimental error, they might well be claimed as having come from a perfect hierarchy by the criteria in vogue. Another hierarchy, chosen at random from those I have formed in the above manner, is as follows :—

—	0·57	0·51	0·46	0·49	0·47	0·39	0·35	0·32	0·26
0·57	—	0·49	0·43	0·51	0·45	0·42	0·33	0·30	0·25
0·51	0·49	—	0·47	0·40	0·41	0·30	0·30	0·27	0·27
0·46	0·43	0·47	—	0·27	0·37	0·18	0·32	0·29	0·38
0·49	0·51	0·40	0·27	—	0·39	0·43	0·22	0·21	0·21
0·47	0·45	0·41	0·37	0·39	—	0·26	0·08	0·28	0·21
0·39	0·42	0·30	0·18	0·43	0·26	—	0·23	0·21	0·15
0·35	0·33	0·30	0·32	0·22	0·08	0·23	—	0·18	0·18
0·32	0·30	0·27	0·29	0·21	0·28	0·21	0·18	—	0·11
0·26	0·25	0·27	0·38	0·21	0·21	0·15	0·18	0·11	—

Sometimes the hierarchical order is almost perfect, as in this example, which happens to be the first I made in this way :—

—	0·77	0·65	0·62	0·58	0·48	0·46	0·33	0·32	0·25
0·77	—	0·59	0·57	0·49	0·51	0·46	0·28	0·20	0·28
0·65	0·59	—	0·52	0·46	0·43	0·39	0·28	0·23	0·23
0·62	0·57	0·52	—	0·44	0·41	0·37	0·29	0·22	0·23
0·58	0·49	0·46	0·44	—	0·33	0·35	0·26	0·19	0·13
0·48	0·51	0·43	0·41	0·33	—	0·26	0·14	0·23	0·18
0·46	0·46	0·39	0·37	0·35	0·26	—	0·14	0·17	0·18
0·33	0·28	0·28	0·29	0·26	0·14	0·14	—	0·13	0·07
0·32	0·20	0·23	0·22	0·19	0·23	0·17	0·13	—	0·00
0·25	0·28	0·23	0·23	0·13	0·18	0·18	0·07	0·00	—

For comparison with these artificial hierarchies may be given a typical hierarchy from experimental psychology.*

---

* Wyatt, 'Brit. Journal of Psychology,' vol. 6, p. 131 (1913).

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}
x_1	—	0·85	0·65	0·67	0·63	0·70	0·61	0·63	0·74	0·64	0·50	0·43	0·54	0·40	0·28
x_2	0·85	—	0·70	0·75	0·62	0·72	0·60	0·71	0·69	0·61	0·41	0·47	0·43	0·33	0·03
x_3	0·65	0·70	—	0·77	0·63	0·66	0·72	0·52	0·47	0·57	0·30	0·56	0·49	0·15	0·006
x_4	0·67	0·75	0·77	—	0·59	0·51	0·57	0·67	0·56	0·52	0·36	0·43	0·46	0·24	0·09
x_5	0·63	0·62	0·63	0·59	—	0·61	0·63	0·50	0·38	0·48	0·53	0·65	0·41	0·32	0·16
x_6	0·70	0·72	0·66	0·51	0·61	—	0·63	0·34	0·43	0·71	0·53	0·34	0·40	0·33	0·15
x_7	0·61	0·60	0·72	0·57	0·63	0·63	—	0·60	0·41	0·38	0·54	0·40	0·46	0·42	0·03
x_8	0·63	0·71	0·52	0·67	0·50	0·34	0·60	—	0·53	0·30	0·36	0·30	0·45	0·28	0·14
x_9	0·74	0·69	0·47	0·56	0·38	0·43	0·41	0·53	—	0·41	0·26	0·20	0·37	0·23	0·31
x_{10}	0·64	0·61	0·57	0·52	0·48	0·71	0·38	0·30	0·41	—	0·26	0·29	0·06	0·29	0·25
x_{11}	0·50	0·41	0·30	0·36	0·53	0·53	0·54	0·36	0·26	0·26	—	0·28	0·26	0·54	0·18
x_{12}	0·43	0·47	0·56	0·43	0·65	0·34	0·40	0·30	0·20	0·29	0·28	—	0·57	0·17	0·11
x_{13}	0·54	0·43	0·49	0·46	0·41	0·40	0·46	0·45	0·37	0·06	0·26	0·57	—	0·20	0·07
x_{14}	0·40	0·33	0·15	0·24	0·32	0·33	0·42	0·28	0·23	0·29	0·54	0·17	0·20	—	0·32
x_{15}	0·28	0·03	0·006	0·09	0·16	0·15	0·03	0·14	0·31	0·25	0·18	0·11	0·07	0·32	—

In searching for some general statement and proof of the above phenomenon, I bethought myself of the somewhat similar phenomenon described in the next section, of which I have long been aware, though it has not, as far as I know, been described, excepting inasmuch as it is implicit in the formulæ there used, which are those of Prof. Pearson and Mr. Filon.

Hierarchical Order produced by Sampling the Population.

In an article entitled "On the Probable Errors of Frequency Constants and on the Influence of Random Selection on Variation and Correlation," in the 'Phil. Trans.,' 1898, vol. 191, A, pp. 229–311, Prof. Pearson and Mr. Filon give the following formulæ :—

$$\mathrm{R}_{r_{12}r_{13}} = r_{23} - \frac{r_{12}r_{13}\left(1 + 2\, r_{12}r_{23}r_{31} - r^2{}_{12} - r^2{}_{23} - r^2{}_{31}\right)}{2\left(1 - r^2{}_{12}\right)\left(1 - r^2{}_{13}\right)},$$

$$\mathrm{R}_{r_{12}r_{34}} = \frac{\left\{\begin{array}{l}(r_{13} - r_{12}r_{23})(r_{24} - r_{23}r_{34}) + (r_{14} - r_{13}r_{34})(r_{23} - r_{21}r_{13}) \\ \quad + (r_{13} - r_{14}r_{43})(r_{24} - r_{21}r_{14}) + (r_{14} - r_{12}r_{24})(r_{23} - r_{24}r_{43})\end{array}\right\}}{2(1 - r^2{}_{12})(1 - r^2{}_{34})},$$

so that, as they say, "errors in the correlations of a first organ with a second and a third have a correlation themselves of the first order," and "errors in the correlation of two organs and in the correlation of a second two have only correlation of the second order."

Suppose now that the correlations among a number of variates taken in pairs are really all the same, and positive, and in a sample let the observed value of r_{38} be the highest observed value :—

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	·	·	·
x_1			h					h				
x_2			h					h				
x_3	h	h	—	h	h	H_2	h	H_1	h	·	·	·
x_4			h					h				
x_5			h					h				
x_6			H_2			—		H_3				
x_7			h					h				
x_8	h	h	H_1	h	h	H_3	h	—	h	·	·	·
x_9			h					h				
x_{10}			·					·				
·			·					·				

H_1 = highest correlation. H_2 = second highest, and so on. h = tendency to be high.

Then, because of the above theorems of Pearson and Filon, the rows and columns x_3 and x_8 will probably contain more total correlation than do the others, and the second highest correlation will probably be in one of these. Let it be r_{36}. Then, after the rows x_3 and x_8 the row x_6 will probably contain many high correlations and r_{86} will probably be the third highest coefficient, because it is a *node* where two *ridges* of high correlation cross. If it is, then the hierarchy so far is excellent, as can be seen on rearranging the square table so as to bring x_3, x_8, and x_6 to the head.

	x_3	x_8	x_6	x_1	x_2	x_4	x_5	x_7	·	·	·
x_3	—	H_1	H_2	h_4	h	h	h	h	·	·	·
x_8	H_1	—	H_3	h_5	h	h	h	h	·	·	·
x_6	H_2	H_3	—	h_6	h	h	h	h	·	·	·
x_1	h_4	h_5	h_6								
x_2	h	h	h								
x_4	h	h	h								
x_5	h	h	h								
·	·	·	·								
·	·	·	·								

Where now is the fourth highest r likely to be found? Somewhere among those marked h for high, and slightly more probably among the r's of the row x_3. If to avoid further rearranging we take it to be r_{31}, then r_{81} is most probably the fifth, and r_{61} the sixth, because they are nodes. And so on. In fact it is clear that when the r's are really equal to one another, then sampling the population will give a set of observed r's which are arranged not in haphazard, but in hierarchical order.

Hierarchical Order produced by Sampling the Elements which make up the Variates.

In the experiment described above with cards and dice, where hierarchical tendency was found to be produced by group and specific factors without any general factor, there was, however, no question of sampling the *population.* The hierarchical order already appears in the theoretical coefficients. There is here, however, another kind of sampling present, viz., sampling of the *elements* which make up the variates.

Let us suppose, instead of deciding the numbers of groups and specific elements as we did by drawing cards, we had dipped into an infinite bag containing black balls and white balls, the former representing group factors and the latter specific factors. Then the most probable event would have been that the proportion of group factors and specific factors in each variate would have been the same as the proportion in which the balls occurred in the bag. If, in addition, we assume the samples drawn to be all the same size, then the most probable result of the whole experiment would have been obtaining all the correlations equal.

That they do not come out equal may be regarded as due to the sampling. The samples vary in size, the proportion of group factors varies from sample to sample, and the distribution of the individual group factors among the several variates departs from the most probable distribution. From this point of view the departure of the correlation coefficients from equality is due to errors of sampling, and the application of the theorem of Pearson and Filon would lead us to expect what as we have seen does actually occur, namely, hierarchical order.

The experiment with the bag of balls would as a matter of fact not produce as great a departure from equality of correlation coefficients as is found in practice in experimental psychology, or as was found in our form of the experiment with card-drawing. This in the latter case is because the cards give a greater variation of the proportion of group to specific factors than would be found with the bag of balls. And this is also the case in mental tests, which are not *chance* samples of the mental elements, but are carefully chosen so as to measure different kinds of activity.

This application of Pearson and Filon's theorem (which contemplates, I take it, only sampling errors in the ordinary sense) to the changes in correlation produced by sampling the underlying elements, is no doubt somewhat bold, and further investigation of an exact analytical nature will probably be fruitful. Several difficulties remain, among which may be mentioned the fact that in sampling the population in the case in question the correlations tend to err all in one way, either rising or falling, whereas here some seem to rise and some to fall, as though the "errors" are the departures, not from the equality point, but from some other origin.

In any case, the analogy between the experiments described and the theory that mental tests are samples of some kind of mental elements is striking and suggestive; and there is evidently no need to postulate a general factor, as has frequently been done. The hypothesis of a general factor is *sufficient* to explain the facts, but is *not necessary*, for a wider theory does so.

An Experimental Determination of the Ionisation Potential for Electrons in Helium.

By Frank Horton, Sc.D., Professor of Physics in the University of London, and Ann Catherine Davies, M.Sc., Royal Holloway College, Englefield Green.

(Communicated by Mr. C. T. R. Wilson, F.R.S. Received December 21, 1918.)

Bohr has developed a theory of atomic structure by applying a system of mechanics based on the quantum theory of radiation to the theory of the nucleus atom proposed by Rutherford. In the case of certain simple atoms he has calculated the minimum voltages through which an electron must fall in order to be able to produce ionisation by collision,* and a comparison of the calculated values with those obtained experimentally serves as a check on the validity of the theory. Next to the hydrogen atom, that of helium has the simplest constitution, and this, for the normal atom, consists of two electrons rotating in non-radiating orbits round a doubly charged positive nucleus. The application of Bohr's formula to helium indicates that the minimum velocity necessary to remove one electron from the helium atom is about 29 volts, but the direct experimental determinations of several observers

---

* N. Bohr., 'Phil. Mag.,' vol. 26, p. 476 (1913).